United Nations DP/FPA/2010/19

Executive Board of the
United Nations Development
Programme and of the
United Nations Population Fund

Distr.: General
16 April 2010

Original: English

Annual session 2010
21 June to 2 July 2010, Geneva
Item 17 of the provisional agenda
UNFPA - Evaluation

UNITED NATIONS POPULATION FUND

BIENNIAL REPORT ON EVALUATION

Report of the Director, Division for Oversight Services

CONTENTS

Page

I. INTRODUCTION 2

II. UNFPA AND PARTNERSHIP FOR DEVELOPMENT 2

III. EVIDENCE-BASED PROGRAMMING 4

IV. UNITED NATIONS REFORM, UNFPA AND THE HARMONIZATION OF EVALUATION METHODOLOGIES 6

V. CAPACITY DEVELOPMENT 7

VI. INDEPENDENCE AND PARTNERSHIP 8

VII. QUALITY AND RELIABILITY OF EVALUATIONS 10

VIII. LEARNING FROM EVALUATION 12

IX. CONCLUSION 13

X. RECOMMENDATION 14

10-31745 (E) 100510

I. INTRODUCTION

1. Since UNFPA presented the periodic report on evaluation (DP/FPA/2008/10) to the Executive Board at its annual session 2008, the adoption of two General Assembly resolutions (62/208 and 63/311), the endorsement of the Accra Agenda for Action at the Third High-Level Forum on Aid Effectiveness in September 2008, and the adoption of Executive Board decision 2009/18 on the UNFPA evaluation policy (DP/FPA/2009/4) have reshaped the process of evaluation. They have redefined the role of evaluation from the objective assessment of an entity's contributions to development results to that of a joint undertaking of development partners with national ownership and leadership. In the light of this reorientation of evaluation, this report: (a) discusses the role of evaluation within the development environment in which UNFPA works; (b) reviews the performance of the evaluation function in UNFPA, especially in the context of United Nations reform; and (c) identifies areas for improvement. This report is timely because of the forthcoming submission to the Board of more than 40 country programme documents that will be developed under the revised United Nations Development Assistance Framework process.

II. UNFPA AND PARTNERSHIP FOR DEVELOPMENT

2. General Assembly resolution 62/208 stresses the "leadership and ownership of programme countries", and General Assembly resolution 63/311 invites the United Nations Secretary-General to propose "an independent, system-wide evaluation mechanism to assess system-wide efficiency, effectiveness and performance". Ownership, the central theme of the Paris Declaration on Aid Effectiveness and the Accra Agenda for Action, is not about technical issues and evaluation methodologies; rather, it refers to the political commitment of all parties involved in development. The emphasis on the leadership and ownership of programme countries has created a de facto shared responsibility and shared accountability to their respective constituencies. This shared responsibility, elaborated in paragraph 24 of the Accra Agenda for Action, is the necessary condition to fulfil the goals of ownership and partnership. The purpose of evaluation is no longer the assessment of results by individual actors in development, but the assessment of their contributions to partnerships, conducted in close collaboration with all stakeholders, implementing partners and beneficiaries.

3. Despite its wide use, the notion of partnership has become increasingly complex. Since the 2002 International Conference on Financing for Development, the concept of development assistance has undergone a revolution, with new goals, an increasing number of partners, and multiple sources of investments[1], as indicated below:

(a) Development assistance targets a multifaceted set of goals. This diversification is the logical consequence of General Assembly resolution 42/183, which endorsed the concept of sustainable development. Beyond its strong traditional focus on accelerating the economic convergence of developing nations with industrialized economies, development assistance is expanding its scope, reflecting the willingness of the international community (donors, programme countries and civil society) to focus also on providing basic human welfare (for example, the Millennium Development Goals) and other public goods[2];

[1] Severino (J.M.) and Ray (O.), "The End of ODA: Death and Rebirth of a Global Public Policy", Center for Global Development, working paper, Washington, DC, March 2009, http://www.cgdev.org/content/general/detail/1421419/.

[2] Public goods are defined here as "a good that is non-rivalrous and non-excludable. This means that consumption of the good by one individual does not reduce availability of the good for consumption by others, and that no one can

(b) Development assistance has become the concern of numerous governmental and non-governmental actors. According to the Organisation for Economic Co-operation and Development (OECD), the annual contribution of non-governmental organizations, including funds, foundations and initiatives, amounts to the equivalent of 15 per cent of the 2008 total official development assistance;

(c) In many cases, the distribution of investments in the production and maintenance of public goods has shifted from external to internal sources. For instance, many middle-income countries invest more in global public goods for health[3] from their own resources than they receive from external resources. This change has placed the decision-making power and initiative in the hands of programme countries. Even though its annual expenditure is small when compared with the $121 billion in official development assistance in 2008, UNFPA has an important role to play, not as a funding agency, but as a facilitator for innovative interventions, for locally or regionally available expertise and capacity, and for community mobilization.

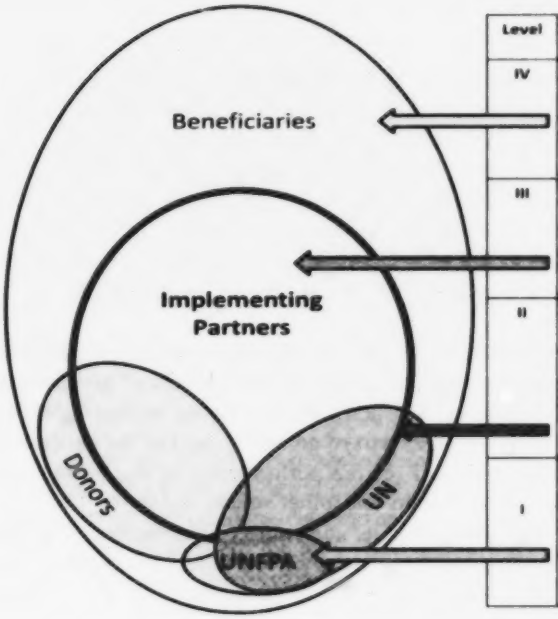
4. In this context of multiple partnerships and programme country leadership, evaluation should be seen as a series of nested functions encapsulated within one another, with different levels of questions to be answered at each level. Levels I to III should address: (a) performance issues within UNFPA; (b) performance issues within the United Nations country team, as part of an assessment of the process of 'delivering as one'[4] and in coordination with other donors; and (c) the performance of the partnership between implementing partners, the United Nations country team and other relevant donors. Level IV should address more complex evaluation questions, and assess results and impact as experienced by the beneficiaries of the partnership between the implementing partners (usually government institutions) and donors. Level IV requires the active participation of actors at the country or community level. These nested evaluation functions are illustrated in figure 1.

be effectively excluded from using the good." See also Joseph E. Stiglitz, *Knowledge as a Global Public Good in Global Public Goods*, ISBN 978-0-19-513052-2, and Stiglitz (J.E.), "The Theory of Local Public Goods Twenty-Five Years After Tiebout: A Perspective", Working paper 954, National Bureau of Economic Research, August 1982.

[3] Woodward (D.) and Smith (R.D.), *Global Public Goods and Health: Concepts and Issues*, WHO, Geneva, 2010.

[4] General Assembly resolution 63/311, paragraphs 18 and 19.

Figure 1: Nested evaluations



Level	Evaluand	Scope
IV	Beneficiaries	Results and impact
III	Partnership between implementing partners, the United Nations and donors	Results, efficiency, effectiveness, relevance, compliance and evaluability Operations research
II	Partnership between implementing partners and the United Nations country team (UNDAF evaluation)	UNDAF - efficiency, effectiveness, relevance, and evaluability Coordination between the United Nations and other donors
I	Partnership between implementing partners and UNFPA (UNFPA country programme evaluation)	Efficiency, effectiveness, relevance, compliance and evaluability Operations research

III. EVIDENCE-BASED PROGRAMMING

5. The expanding array of contributors to development parallels an expanding array of instruments. While microfinance organizations and community development banks have flourished, other tools using market mechanisms, such as microcredit[5] community development venture capital or patient capital[6], have also been successful. Yet, despite this evolution, official development assistance remains measured, either by financial volume or by politically defined targets, such as the Millennium Development Goals. As pointed out by several authors[7], while measuring inputs is useful, there is a need to focus more on measuring results and impact. Donors (including bilateral and multilateral organizations, non-governmental organizations and foundations) typically complete internal evaluations that focus on individual institutions and do not capture the overall impact of policies and programmes at country, regional or global levels.

[5] Banerjeey (A.), Duflo (E.), Glennersterx (R.), Kinnan (C.), "The Miracle of Microfinance? Evidence from a Randomized Evaluation", MIT, May 30, 2009.

[6] Patient capital is a third way that seeks to bridge the gap between the efficiency and scale of market-based approaches and the social impact of pure philanthropy. See Novogratz (J.), "The Blue Sweater: Bridging the Gap Between Rich and Poor in an Interconnected World", 2009, MacMillan, USA.

[7] See note 3 and Michael A. Clemens and Todd J. Moss (2007), "The Ghost of 0.7%: Origins and Relevance of the International Aid Target", International Journal of Development Issues 6 (1): 3-25.

6. Within the evaluation community, a consensus exists that the quality of evaluation is most closely linked to the quality of programme design. Good-quality programmes are more likely to achieve desired results and impact, and to be amenable to evaluations that can verify the results and identify elements of how the programme was implemented, as a basis for lessons learned. This simple assertion implies a more complex debate about the definition of what an effective programme is. For the purpose of this report, one may reflect the general consensus that effective programmes are theory-based, offering a conceptual rationale, in addition to their achieving desired outcomes or outputs. Evidence can be generated by macro-level organizational change theory or systems analysis, which is based on a thorough and systematic data review[8] and by operations research[9]. Both approaches are not mutually exclusive; their combination is one of the best approaches to establish a fruitful relationship between policy and research.

7. During the 2008-2009 period covered by this report, the UNFPA Division for Oversight Services developed and implemented a standardized method to review UNFPA country programme design. The method focuses on three elements: (a) relevance; (b) strategic alignment; and (c) the quality of monitoring and evaluation. Based on the results of oversight work in 2008 and 2009 that included the review of 11 country programmes, several recurring themes emerged.

8. With regard to relevance, programme design in country programmes deemed partially satisfactory or unsatisfactory lacked evidence-based information. Such information is necessary to support the choice of UNFPA intervention strategies and geographical subregions. Though programme design tries to meet multiple needs, it would be preferable to focus on a limited set of interventions with demonstrated evidence of effectiveness. This would avoid dispersed and small-scale efforts managed through a large number of annual work plans, with significant administrative overhead.

9. With regard to strategic alignment, UNFPA country programmes are generally well-aligned with national development frameworks. However, in some country programmes, the pivotal role of UNFPA with regard to the implementation of the census and capacity development to support the use of data is absent from the United Nations Development Assistance Framework. This raises the risk that these activities might not benefit from the support of the United Nations country team, despite the widely agreed benefits of their outputs (for example, census data and the improved use of data in planning) for all development partners.

10. With regard to monitoring and evaluation, country programmes deemed partially satisfactory or unsatisfactory usually lack indicators, and indicators lack baseline data and/or targets and the means of verifying progress. Monitoring and evaluation must evolve in UNFPA so that the necessary indicators, baseline data and/or targets are developed during the programme design phase and integrated into the programme. Table 1, below, provides an overview of recurrent findings of the Division for Oversight Services.

[8] Caldwell J.C., Phillips J.F., Khuda B. (eds). 2002. Special Issue: Family planning programs in the twenty-first century. Studies in Family Planning 33 (1).

[9] Carlos Brambila, Emma Ottolenghi, Celeste Marin and Jane T. Bertrand, "Getting results used: evidence from reproductive health programmatic research in Guatemala". Health Policy and Planning 2007; 22:234–245.

Table 1: Evaluation of country programmes: recurrent findings

Domain	Recurrent weaknesses/high risks	Potential remediation and mitigation measures
Relevance	Lack of focus on interventions with evidence of effectiveness; lack of an evidence-informed assessment of needs	Structured technical review of planned interventions during programme development
Strategic alignment	Absence of census from United Nations Development Assistance Framework and United Nations country team workplans	Systematic regional office review to ensure inclusion of census in United Nations Development Assistance Framework and mobilization of United Nations country team support for census
Quality of monitoring and evaluation	Missing indicators, and/or baselines and/or targets, and no evidence of using knowledge and monitoring information systematically	Referral or conditional approval of country programme documents submitted without results frameworks that include specified indicators, baselines and targets

11. The review by the Division for Oversight Services could neither indicate a significant investment in operations research nor could it demonstrate a reliable evaluability of programmes that would provide information on impact or results. This would enable knowledge utilization for subsequent programme cycles. Greater efforts by UNFPA are needed to make evidence-based programming possible for all programmes, as requested in paragraph 10 of Executive Board decision 2009/18.

12. By ensuring that evidence-based information and benchmarks are available, UNFPA has the opportunity at the programme-inception phase to: (a) build a solid partnership with implementing partners and stakeholders; (b) elaborate rational programme content where capacity-building is linked to clear priorities, where gender and community mobilization strategies relying on cultural and religious values are interwoven in the programme, and where innovative interventions are tested with appropriate methodologies[10] for scaling up; and (c) demonstrate a positive contribution to the efforts and investments deployed by programme countries to improve the quality of life of beneficiaries.

IV. UNITED NATIONS REFORM, UNFPA AND THE HARMONIZATION OF EVALUATION METHODOLOGIES

13. The methodological issue in evaluation today is less about having the 'right' methodology than about establishing and following methodological guidance that would harmonize approaches within UNFPA as well as within the United Nations system, in order to facilitate comparisons of results. Harmonized methods imply a convergence of evidence-informed programming as a key condition for successful evaluation; they would facilitate the elaboration of programme content with appropriate results-based management and indicators, and would thus lead to programmes with adequate levels of evaluability. However, evaluations are performed using different methodologies, which makes comparisons and benchmarking difficult. The consistent application of methodologies is necessary in order to avoid an 'evaluation gap' that affects many institutions that provide assistance. UNFPA is no exception. The review of evaluation methodologies by the Division for Oversight Services indicates wide

[10] Duflo (E.), Kraemer (M.), "Use of Randomization in the Evaluation of Development Effectiveness", MIT, 2009.

disparities and calls for efforts to harmonize not only within UNFPA but also with other United Nations partners.

14. With the approval by the Executive Board in 2009 of the UNFPA evaluation policy, UNFPA has taken steps to strengthen evaluation and comply with General Assembly resolution 62/208. While the UNFPA policy may strengthen evaluation, it can do little to address the system-wide coherence envisioned in General Assembly resolution 63/311, endorsing a system-wide evaluation function in the absence of a United Nations-wide commitment to harmonization. At present, the main vehicle for such efforts is the United Nations Evaluation Group. In the light of constraints imposed by the United Nations Evaluation Group, the Division for Oversight Services completed a review of all United Nations Evaluation Group working groups and task forces in 2009, to identify which ones would be most relevant for UNFPA. The review prioritized several United Nations Evaluation Group activities in which UNFPA was able to participate in the latter part of 2009. The United Nations Evaluation Group, at its global meeting in April 2010, will discuss the issue of extending participation in United Nations Evaluation Group task forces to those responsible for decentralized evaluations.

15. The harmonization of methodologies is thus a long-term project. Based on the varying capacity for evaluations of United Nations system organizations, any entity designed or constituted to perform a system-wide evaluation function, as per General Assembly resolution 63/311, will likely be more effective if designed or constituted with flexibility, in order to maximize participation and collaboration among the evaluation units of various organizations.

V. CAPACITY DEVELOPMENT

16. In recent years, about one quarter of donor aid, or more than $15 billion a year, has been invested in technical cooperation, the major part of which is ostensibly aimed at capacity development[11]. General Assembly resolution 62/208 pays great attention to capacity. In response, United Nations entities have felt compelled to emphasize their efforts and commitment to developing capacity. The most common misunderstanding about capacity development relates to its conceptualization as a unidirectional knowledge transfer from those who have expertise to those who need it. Evidence indicates that capacity development is unlikely to succeed if the following conditions are not met:

(a) broad-based participation and a locally driven agenda;
(b) building on local capacities;
(c) ongoing learning and adaptation;
(d) long-term investments;
(e) integration of activities at various levels to address complex problems.[12]

17. According to OECD, capacity development should be understood as an "endogenous process, strongly led from within a country, with donors playing a supporting role. According to this vision, political leadership and the prevailing political and governance system are critical factors in creating opportunities and setting limits for capacity development efforts. Country policy ownership is not a simple yes/no issue, but a matter of processes and trends".

[11] OECD, *The Challenge of Capacity Development: Working Towards Good Practice*, OECD Journal on development, Vol. 8, No. 3 – ISSN 1816-8124 – © OECD 2008.
[12] Canadian International Development Agency, "Why, What and How", CIDA, Policy Branch Vol. 1, No. 1, Gatineau, Quebec, Canada, May 2000.

19. Thus, capacity development is no exception to other areas of development. It is a partnership. This partnership requires better understanding among all parties to be able to respond to the widening scope of evaluation, the diversification of tools, and the momentum towards joint evaluations within the rapidly evolving development assistance environment. If evaluation is to be understood in this context, then all evaluation methodologies should explain how the programme interventions implemented and assessed capacity-development efforts. Without this change, it will be difficult to implement paragraph 129 of General Assembly resolution 62/208 on strengthening evaluation capacities in programme countries.

19. UNFPA is making progress on this issue. Regionalization and the development work for the UNFPA evaluation policy during 2008-2009 created an internal focus on UNFPA evaluation capacity development efforts. In October 2009, the Africa Regional Office, with the participation of staff from the Learning and Career Management Branch, Division for Human Resources; the Programme Division; and the Division for Oversight Services held a week-long training programme for country office monitoring and evaluation officers and focal points. UNFPA commitment to developing and supporting regional evaluation capacity will, in turn, strengthen the capacity of programme countries to undertake evaluations. Thus, with regard to capacity development, it is too soon to assess progress since the approval of the evaluation policy in June 2009, and to provide a definitive assurance statement on UNFPA activities undertaken to enhance the evaluation capacity of its implementing partners.

VI. INDEPENDENCE AND PARTNERSHIP

20. The shared responsibility of governments extends to evaluation, focusing on results or on the contribution to the results of development partnerships. The notion of shared accountability calls for a re-examination of the concept of independence in evaluation. The focus of General Assembly resolution 62/208 on joint efforts and national ownership has created tension about the issue of independence and uncertainty about the value and risks of evaluation partnerships; in many cases, it has created institutional worries in the United Nations system.

21. The dilemma can be described as follows. On the one hand, the consensus on partnership in evaluation and the recognition of 'leadership and ownership of programme countries in the evaluation of all forms of assistance, including that provided by the United Nations development system', implies that evaluation occurs where assistance takes place and by the actors providing and receiving the assistance. Evaluation is thus a decentralized process that is embedded in development activities where they are implemented; this process legitimizes the evaluation of donor programmes by programme countries. On the other hand, as long as evaluation capacity has not reached the desired level in programme countries, the risk exists that professionals undertaking evaluation could also be involved in implementing programmes, thus compromising independence and objectivity. This dilemma can be avoided by adopting different modalities that guarantee independence for each level of evaluation, as described in figure 1 (nested evaluations). Table 2, below, which proposes mechanisms to guarantee the independence of evaluators, illustrates that reliance on and compliance with globally recognized evaluation norms and standards may be the most effective way of guaranteeing independence, even in the context of complex partnerships.

Table 2 Evaluations and independency mechanisms

Level	Evaluand	Scope	Mechanisms to guarantee independence
IV	Beneficiaries, partnerships, implementing partners and donors	Results and impact	Third party evaluations, globally accepted norms and standards, transparency (for example, disclosure of programme theory, methodology, data, analysis and results), ownership and leadership by implementing partners
III	Partnerships, implementing partners and donors	Results, efficiency, effectiveness, relevance, compliance and evaluability Operations research	Third party evaluations, globally accepted norms and standards, transparency (for example, disclosure of programme theory, methodology, data, analysis and results) Ownership and leadership by implementing partners
II	United Nations country team United Nations reforms	United Nations Development Assistance Framework Efficiency, effectiveness, relevance, evaluability Coordination between the United Nations and other donors	United Nations evaluation team or third party, assurance provided by oversight units on design, methodology, collection and analysis of data, results and report, globally accepted norms and standards, transparency (for example, disclosure of programme theory, methodology, data, analysis and results), independent system-wide evaluation mechanism (General Assembly resolution 48/162)
I	UNFPA country programmes	Efficiency, effectiveness, relevance, compliance, evaluability Operations research	UNFPA or third party, assurance provided by oversight units on design, methodology, collection and analysis of data, results and report, globally accepted norms and standards, transparency (for example, disclosure of programme theory, methodology, data, analysis and results)

22. Since the UNFPA policy on evaluation has only recently been adopted, evaluations conducted by UNFPA country offices do not yet reflect the revised approach. All evaluations conducted between 2009 and February 2010 are level 1 evaluations, with little or no participation of implementing partners and beneficiaries. None of the evaluations address the UNFPA contribution to the embedded partnerships, even in countries with programme-based or sector-wide approaches.

23. To implement the evaluation policy, UNFPA and the Division for Oversight Services clarified roles and responsibilities with regard to planning, implementation, supervision, quality assurance and follow-up of evaluations. The Programme Division is responsible for organizational learning, supporting

country programme evaluations and follow-up of evaluations, the Division for Oversight Services is responsible for quality assurance and independent thematic evaluations.

28. The UNDP's evaluation policy distinguishes between centralized and decentralized evaluations. The Division for Oversight Services carries out centralized evaluations with the primary purpose of assurance, while management undertakes decentralized evaluations to build the evidence base of results achieved and to identify lessons learned from effective programming. With the support of the regional offices, the Programme Division facilitates fund-wide coordination.

29. In contrast with other United Nations organizations, UNDP has evaluation resources throughout the organization rather than in one organizational unit. Since more than one division is responsible for evaluation, UNDP has, in 2009 and 2010, strengthened these divisions and a number of country offices in terms of human resources, in order to respond to the Executive Board and to implement the evaluation policy. One challenge for UNDP is, especially at the country level, is that some staff members must assume the additional functions of serving as focal points on monitoring and evaluation, which will increase their already heavy workload. This makes it difficult to move as quickly as needed to implement the evaluation policy.

30. Based on the evaluation policy, UNDP's would have adequate staff to conduct evaluations if all such staff had sufficient expertise. This is a critical point for the debate on independence. The quality of design, methodologies, implementation and conduct of evaluations is a prerequisite for independence, an area that needs greater attention by UNDP's.

VII. QUALITY AND RELIABILITY OF EVALUATIONS

31. The international community can draw on a variety of publications on evaluation quality and reliability, including norms and standards. UNDP's relies on United Nations Evaluation Group standards. These norms and standards are in line with best practices sanctioned by professional evaluation associations around the world. The starting point of a global partnership on evaluation should thus be a global consensus and partnership on norms and standards, a necessary condition for ownership and compliance by all involved parties, including civil society organizations and beneficiaries of development interventions.[7] Accordingly, UNDP's should comply with its evaluation policy and meet quality standards in its evaluative work. To assess the degree of compliance, the Division for Oversight Services conducted an evaluation quality assessment exercise in 2009 that included evaluations performed during the period covered by this report (2008-2009). The results of this review are available on www.undp.org/public oversight and are summarized below.

32. All UNDP's business units report their evaluation work through an annual reporting process. For 2008, country offices reported 102 evaluations. For 2009, country offices reported 139 evaluations. In addition, the Division for Oversight Services completed three thematic evaluations (two in 2008 and one in 2009), and the evaluation quality assessment.

33. Thematic evaluations of UNDP's experience in assistance effectiveness, through participation in sector-wide approaches and joint programming, yielded findings in the latter half of 2008. Such evaluations noted that UNDP's had been among the leading United Nations organizations to engage in

[7] Chianca, T. (). "The OECD/DAC Criteria for International Development Evaluations: An Assessment and Ideas for Improvement". Journal of Multidisciplinary Evaluation, Vol. 5, No. 9. ISSN 1556-8180. March 2008.

these modalities, and that the coordination workload required to ensure successful implementation of and participation in joint programming and sector-wide approaches was substantial. The 2009 evaluation of the UNFPA humanitarian response provided management with recommendations to improve the progress made since the endorsement by the Executive Board, in decision 2006/35, of the UNFPA strategy for emergency preparedness, humanitarian response, and transition and recovery. The recommendations of the 2009 evaluation focused on mainstreaming humanitarian response in UNFPA country programmes. They also focused on reviewing business processes to increase UNFPA effectiveness in timely procurement and hiring during humanitarian crises.

30. UNFPA requested all country offices, regional offices and headquarters units to submit evaluations completed during 2007-2008, as part of the evaluation quality assessment process. For the period 2007-2008, the Division for Oversight Services received 57 evaluations completed by UNFPA country offices, a significantly lower number than the 197 evaluations reported by country offices for the same period, and the 116 reported for the 2007-2009 period. The Division for Oversight Services received no submissions from regional offices or headquarters units, whereas the former reported having completed four and the latter 11 for the same period. The discrepancy between the number of evaluations that are reported by business units in their annual reports and the number of those provided to the Division for Oversight Services for review is of concern. Either the self-reporting of the entities is inaccurate or the entities are not submitting the documents to the Division for Oversight Services, which is mandated by the UNFPA oversight policy to receive such evaluations. Regardless of the cause, UNFPA must address this deficiency.

31. In the absence of the systematic recording of evaluations by management, the Division for Oversight Services is unable to provide assurance on the number of evaluations completed in UNFPA. Of the 57 evaluations submitted by country offices during the 2007-2008 biennium, 17 were evaluations of UNFPA country programmes, nine were midterm reviews and 31 were evaluations of other programmes or projects supported by UNFPA. As in previous years, teams of national and international evaluators conducted most of the country-level evaluations.

32. The 2009 evaluation quality assessment exercise reviewed 37 of 57 evaluations. Overall, results show little change from the 2008 and 2005 evaluation quality assessment exercises. The large proportion of evaluation reports that lack basic elements of evaluation (such as clear statements of evaluation questions, methods suitable to answer the evaluation questions, findings arising from data and evidence gathered through the application of said methods, and conclusions and recommendations grounded in the findings) suggests that quality could be improved by implementing a systematic review and coaching process for UNFPA evaluations. Modest improvements in the evaluation of effectiveness, the validity of methods and attention to gender mainstreaming are evidence that improvements are possible. As quoted in published reviews[10], this situation is not unique to UNFPA.

33. Attention must be given to evaluating programme relevance and impact. Both dimensions are assessed in UNFPA evaluations as the proportion of evaluations meeting standards in these areas. This proportion has fallen precipitously. The 2009 evaluation quality assessment report provided to management includes recommendations, which, when implemented, would improve the quality of evaluation, based on their success in other settings.

[10] Barham D. Savedoff et al., "When Will We Ever Learn? Improving Lives through Impact Evaluation", Report of the Evaluation Gap Working Group, Center for Global Development, Washington, D.C., May 2006.

34. In order to ensure consistency between the number of evaluations reported by UNFPA entities and the number that were submitted to the evaluation quality assessment process, UNFPA must develop systems to track and verify evaluation activities. This is a necessary step to provide assurance and ensure that lessons learned are disseminated as widely as possible. The quality of the modest number of evaluations available for review demonstrates scope for improvement.

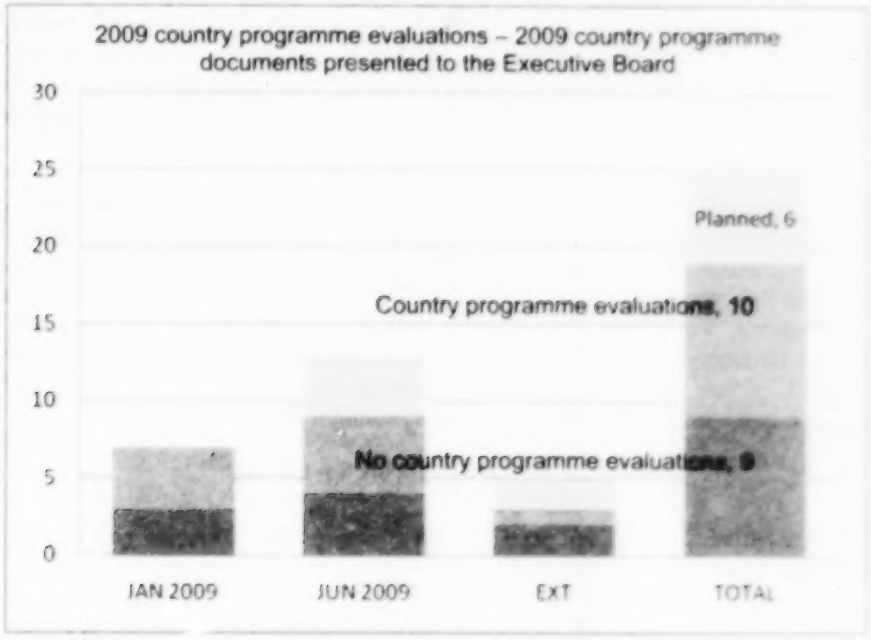
VIII. LEARNING FROM EVALUATION

35. For the purpose of evaluation, transparency refers to at least three criteria: (a) the accuracy and reliability of the evaluation findings, which in turn depends on the level of qualification and independence of the evaluators; (b) the reliability of the evaluation itself (globally accepted norms and standards); and (c) the dissemination and use of the report, including recommendations. The consensus on a wide dissemination of results should be the rule, reflecting the notion that partnership extends to the public at large so that evaluation findings can inform and strengthen not only governmental partners but also civil society partners and stakeholders.

36. UNFPA management has informed the Division for Oversight Services that it is developing systems, for deployment in 2010, to ensure a systematic response to and follow-up of decentralized evaluation recommendations and to disseminate lessons learned. For centralized evaluations, the Division for Oversight Services extended the Comprehensive Assessment Recommendations Database Systems used to track audit recommendations to include recommendations from centralized evaluations completed from 2009 onward. In addition, all evaluations reviewed through the evaluation quality assessment process have been made available to the research and professional communities through an evaluation database developed and managed by the Division for Oversight Services. Paragraph 20 of the UNFPA evaluation policy (DP/FPA/2009/4) states that "in accordance with the UNFPA accountability framework and the UNFPA oversight policy, the Executive Director will normally disclose reports to the public, while exercising discretion in protecting the legitimate rights of programme countries." In line with this policy, the contents of the evaluation database are available on the UNFPA website (www.unfpa.org/public/about/oversight).

37. In the light of the centrality of UNFPA country programmes, the Division for Oversight Services reviewed the ability of UNFPA to plan evaluation in a timely manner and to use country programme planning cycles as an opportunity to influence decisions. Of the country programmes considered by the Executive Board in 2009, 40 per cent had been evaluated so as to inform the design of the subsequent programmes. A further 24 per cent planned to do evaluations; these may be unlikely to inform the development of the subsequent country programme due to the evaluations occurring after Executive Board approval of the programme. The review of evaluation activities in relation to UNFPA country programmes submitted to the Executive Board in 2009 is summarized in figure 2.

Figure 2: Country programme evaluations for 2009 country programme documents



IX. CONCLUSION

38. UNFPA has taken important steps to strengthen evaluation. Accomplishments of the 2008-2009 period include: (a) the development of an evaluation policy; (b) the establishment and staffing of regional monitoring and evaluation posts in all UNFPA regional offices; and (c) the appointment of monitoring and evaluation focal points in a number of country offices.

39. Recurring evaluation findings and recommendations highlight scope for improvement in evidence-based programme design, programmatic focus, and attention to monitoring and evaluation. In addition, an analysis of the coverage of evaluations suggests that UNFPA needs to take action to ensure that it complies with paragraph 8 of Executive Board decision 2009/18, and that it ensures that all country programmes are evaluated at least once during their cycle.

40. In 2009, the Division for Oversight Services assessed the quality of evaluation at UNFPA. Even though there are no major differences from the 2005 evaluation quality assessment, the Division has noted improvements in the assessment of effectiveness, the validity of methods and gender mainstreaming. Attention must be paid to assessing the relevance and impact of UNFPA efforts. Improving quality will require an enhanced culture of evaluation and accountability for results within UNFPA, a finding that is not uncommon in the United Nations system. Capacity development to support such a culture will need to be negotiated with commitments to joint evaluations and strengthened national ownership of evaluation. A systematic management response and follow-up to evaluation recommendations will ensure that UNFPA maximizes the effective use of its resources in achieving its goals.

41. The challenge for UNFPA is to improve its internal evaluation processes while keeping pace with the national emphasis on ownership and leadership. These trends challenge the role of UNFPA and of other United Nations organizations. Such an examination will need to address the nature of the UNFPA role in the three-way relationship between: (a) the accountability requirements of donors attached to development assistance funding; (b) the increased ownership and leadership by programme countries of programmes and their evaluation; and (c) the UNFPA desire to demonstrate progress in achieving its mandate.

42. General Assembly resolution 62/208 can be understood as committing UNFPA and programme country governments to evaluating issues of mutual interest, focusing on the development results to which UNFPA and programme countries are committed. This commitment implies that evaluation will be of most value when it focuses on the results experienced by beneficiaries. For UNFPA, these results would include, for instance, the numbers of pregnant women who no longer die during childbirth, and a reduction in the frequency of gender-based violence. Accepting this as a suitable approach to evaluation, UNFPA should pursue this matter while reviewing its strategic plan, 2008-2013, as well as institutionalize it in the subsequent strategic plan, beginning in 2014.

X. RECOMMENDATION

43. The Executive Board may wish to take note of the biennial report on evaluation, as contained in the present document (DP/FPA/2010/19).